Measurement error of visual casting surface inspection

Gokcer Daricilar

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Recommended Citation

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Measurement error of visual casting surface inspection

by

Gokce Daricilar

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Industrial Engineering

Program of Study Committee:
Frank E. Peters, Major Professor
Matthew C. Frank
Iver E. Anderson

Iowa State University
Ames, Iowa
2005

Copyright © Gokce Daricilar, 2005. All rights reserved.
This is to certify that the master’s thesis of

Gokcer Daricilar

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy
“I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind.”

Lord Kelvin
# TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION 1

CHAPTER 2. LITERATURE REVIEW 3

CHAPTER 3. DEFINING MEASUREMENT ERROR 5

3.1. Defining Repeatability 7
3.2. Defining Reproducibility 10

CHAPTER 4. METHODOLOGY 13

4.1. Image Data Collection 13
4.2. Sticker Coordinate Data Extraction 16
4.3. Statistical Analysis via Automated Algorithms 21

CHAPTER 5. RESULTS 27

CHAPTER 6. DISCUSSION 35

6.1. Effects of Casting Cleaning Procedures 35
6.2. Determination of Search Zone Coefficients 37
6.3. Sensitivity Analysis for Search Zone Coefficients 39
6.4. Implications of Results 41

CHAPTER 7. CONCLUSIONS 43

7.1. Methodology 43
7.2. Steel Casting Industry 43

CHAPTER 8. FUTURE WORK 45

APPENDIX A. INSTRUCTIONS FOR INSPECTION DATA COLLECTION 46
APPENDIX B. CLEANED IMAGE SAMPLES 50
APPENDIX C. DVT FRAMEWORK SCRIPT 53
REFERENCES CITED 57
ACKNOWLEDGEMENTS 59
VITA 60
CHAPTER 1. INTRODUCTION

The issues related to defect formation are extremely important, but as important is the ability to determine the quality requirements of a casting purchaser and supplier. Therefore, visual inspection of the surface quality is an imperative task conducted during the processing of steel castings. These inspections identify the occurrence of unacceptable casting surface defects, such as inclusions, porosity, burnt on sand, and flash. The castings that are marked for surface defects are then taken through a series of cleaning operations, where the marked defects are welded and/or ground to surface quality specifications.

Recent studies show that currently there is no satisfactory method, whereby the surface quality requirements can be communicated throughout the manufacturing and purchasing phases of casting production. Although there are some standards in existence, they are unusable in a number of ways. For example, various standards exist that utilize replicas of cast surfaces, but they do not allow the surface to be inspected to a specific level. The result is that the description of casting surfaces is uncertain at best and impossible at worst. The lack of a reasonable measurement system for quality causes several implications, including uncontrolled processing times. Undetected surface defects during the visual inspection process will result in unacceptable quality standards and returns from the customer. Marking minor surface imperfections as defects will result in excessive rework.

The current visual inspection procedures used have never been subjected to a statistical study to determine their usefulness to the industry. Studies on other quality measurements such as radiography have shown their inability to characterize even the most popularly used quality standards. For example, a study by Berdonosov, Kuleshov, and Vyatkino (1992) dealt with the problems of increasing the defect detection ability using
nondestructive radiographic techniques. Halmshaw (1999) questioned the reliability of radiographic methods in a recent study. Another study by Anon (1996) showed how the ability of an inspection system to discriminate between defective and sound material is decreased by increased inspection sensitivity in the inspection process of high quality castings.

This study looks at the problems with the current visual surface inspection standards used in the casting industry today. The goal of this study is to assess the amount of variation introduced by these inspections to determine the measurement error associated with this process. The objective is to develop a methodology to quantify the amount of variation in terms of repeatability (variation within the same operator) and reproducibility (variation between different operators) and apply it to the visual assessment data collected at three different steel foundries to draw beneficial conclusions.
CHAPTER 2. LITERATURE REVIEW

A review of variability detection studies conducted in the areas of surface inspection and surface measurement were analyzed as a starting point for this study. Although no studies in the visual assessment of casting surface inspections were found, many studies in the optical casting surface defect detection, recognition, and classification were noted. A very similar gage repeatability and reproducibility (gage R&R) study in subjective evaluation of image data was also carefully examined to generate valuable insight.

A study by Wong, Elliot, and Rapley (1995) on surface defect detection with high integrity castings admitted little has been reported in the previous studies on recognition and classification of the casting defects. The study utilized a series of fuzzy logic based algorithms from which very encouraging results were obtained.

Another study with encouraging results looked at an automated visual inspection system for the detection of defects in a range of images of castings (Newman et al., 1995). The system used computer-aided design (CAD) model information and inspection algorithms in several stages, including surface classification and inspection. The inspection system was able to correctly classify over 90% of the casting images.

Some not so successful attempts were also made in the automated detection of surface defects in machined castings. A study by Woods and Allen (1989) utilized computer vision to automate the inspection of machined steel components for cracks using the fluorescent magnetic particle method. Unfortunately, the study's results indicated that although the candidate generation stage was promising, reliable detection of all defects was achieved only at the cost of an unacceptably high overall false positive rate.
Although some success was achieved in the automated surface inspection systems, they were limited to a specific type of surface defect, such as surface cracks. These automated systems are usually not practical, as the castings need to be inspected for a number of different types of surface defects. A study by Someji et al. (1997) explained that because of the various defect types and their complicated shapes, defect inspection of castings is dependent on the human eye. However, the study also claimed that this process could get unstable by the tiredness of a person or a change in the environment, and become highly subjective.

Another study in the steelmaking process by Reisinger and Kogler (2001) attempted to combine the human methods with automated visual surface inspection systems. They claimed that in order to obtain excellent inspection results and, especially, detect smaller defect sizes, steel strip inspection had to be carried out by both human and an automatic inspection system. The study had proved successful in meeting the specific inspection requirements with this combined methodology.

A gage R&R study by Lee et al. (1992) on subjective evaluation of image data from the medical industry was also examined. The study used nine medical x-ray CT head images from three patients as test cases. Six radiologists participated in reading the 99 images (some were duplicates) compressed at four different compression ratios, original, 5:1, 10:1, and 15:1. The study found that the six readers agreed more than by chance alone and their agreement was statistically significant, but there were large variations among readers as well as within a reader. What this study analyzed is very similar to the visual casting surface inspections, as both are highly subjective because of their reliance on the human eye.
CHAPTER 3. DEFINING MEASUREMENT ERROR

As mentioned in Chapter 1, defect detection is a very important task that is dependent on the human eye during the casting surface inspection process. In order to assess the amount of variation introduced by the visual casting surface inspections, the measurement error associated with this process has to be defined. The definition of the measurement error must be able to quantify the amount of variation between different inspection trials of the same casting in terms of repeatability and reproducibility.

Visual surface inspections of the castings are performed by operators who identify certain areas of the casting that need grinding or welding as defects, and mark those areas with a marker. This study utilized inspectors that perform these tasks on a daily basis. This required a total of four inspection trials for each casting, as every casting needed to go through two operators twice. After the castings were marked for defects, round stickers of specific size (depending on the size of the casting) were used to cover the markings as shown in Figure 1. This made locating and quantifying the size of the marked areas possible. More detail about the sticker size and the application process can be found in Chapter 4.

![Casting Defects](image1)
![Inspection Markings](image2)
![Sticker Application](image3)

Figure 1. Example of the marked casting defects covered by the stickers.
In this study, a cluster of stickers in the same region is called a master cluster and defined as a group of stickers that have contact with each other and located anywhere among the four combined inspection trials of the same casting. Master clusters represent the supposed defect regions identified and marked by an operator during the inspection trial of a casting. Master clusters are determined with the use of a user defined search zone coefficient. The search zone coefficient is a number, which, when multiplied by the radius of a sticker, creates a circular search zone around the sticker's center position. During the master cluster determination process, if the center position of another sticker from any of the four inspection trials of the same casting falls within this zone, then the two stickers are considered to be touching and assigned to the same master cluster region. It is also possible for a sticker to have multiple touching stickers if the center positions of more than one sticker fall within the search zone, especially from different inspection trials. More detail on the determination of search zone coefficients and a sensitivity analysis are included in Chapter 6.

The master cluster concept is introduced to characterize a marked area by an operator as a supposed defect region. A visual representation of the master cluster concept is displayed in Figure 2. In this case, operator 1's trial 1 contains 7 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. On the other hand, operator 2's trial 1 contains 5 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. Therefore, the four combined inspection trials of the same casting by two different operators result in 4 master cluster regions. The combining of the inspection trials of a casting can also be described as superimposing its inspection images on top of each other. This combination operation is displayed with a union (\(\cup\)) symbol in the following figures.
Figure 2. Defining master clusters with the four combined inspection trials of the same casting. The two stickers at the bottom of the casting from trial 2 of operator 1 both fall in the same master cluster region, since they are connected by stickers from other trials.

3.1. Defining Repeatability

Repeatability error of an operator is the amount of variation between the two inspections of that operator of the same casting. In this study, repeatability is defined by two aspects. The first aspect represents how well an operator performs in identifying the same supposed defect regions between the two inspection trials of the same casting. The other aspect represents how well the operator performs in defining the size of the supposed defect regions between the two inspection trials of the same casting.

After a casting goes through an operator twice and the markings of the operator are covered with stickers by the trial moderator, the two inspection trials are compared by one-to-one sticker matching to determine the repeatability of that operator. This one-to-one
sticker matching process is also performed with the user defined search zone coefficient. During the one-to-one sticker matching process, if the center position of another sticker from the other inspection trial of the same casting falls within this zone, then the two stickers are considered to be matched. It is also possible for a sticker to have multiple matching stickers if the center positions of more than one sticker fall within the search zone.

The repeatability of an operator is reported as two different percentage values. The first percentage value indicates how well the operator performed in identifying the same supposed defect regions between the two inspection trials of the same casting, reported as percent master cluster match. The second percentage value indicates how well the operator performed in defining the size of the supposed defect regions between the two inspection trials of the same casting, reported as percent sticker match. The higher the master cluster or sticker percentage values, the better their corresponding repeatability measurements are. For example, a 0% match would be the lowest, indicating no repeatability, where a 100% match would be the highest, indicating perfect repeatability for an operator’s performance on a particular casting.

A visual representation for the repeatability of two different operators can be seen in the following figures. In the case shown in Figure 3, operator 1’s trial 1 contains 7 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. The two inspection trials of the same casting are then compared. This comparing operation is displayed with an intersection (\(\cap\)) symbol in the following figures. Operator 1’s two inspection trials for the same casting match 7 stickers out of 11, and 2 master clusters out of 3 in total. For this hypothetical example, operator 1’s repeatability is 64% for sticker and 67% for master cluster matching, respectively.
Figure 3. Defining repeatability of operator 1 with the same casting.

In the case shown in Figure 4, operator 2’s trial 1 contains 5 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. The two inspection trials of the same casting are then compared. Operator 2’s two inspection trials for the same casting match 2 stickers out of 9, and 1 master cluster out of 4 in total. For this example, operator 2’s repeatability is 22% for sticker and 25% for master cluster matching, respectively.

Between the two examples shown, operator 1 displays better repeatability with higher percentage values for both sticker and master cluster matching. Therefore, it can be concluded that operator 1 performed better than operator 2 in both identifying the same supposed defect regions and using the same size markings to define the same supposed defect regions between the two inspection trials of the same casting.
3.2. Defining Reproducibility

Reproducibility error between operators is the amount of variation between the inspections of those operators of the same casting. In this study, reproducibility is also defined by two aspects. The first aspect represents how well two or more operators perform in identifying the same supposed defect regions between the two inspection trials of the same casting. The other aspect represents how well the operators perform in defining the size of the supposed defect regions between the two inspection trials of the same casting.

After a casting goes through both operators twice and the markings of the operators are covered with stickers by the trial moderator, the two inspection trials of the operators are combined to represent the best evaluation of each operator for that casting. Therefore, better repeatability from the operators will result in higher reproducibility between the operators.
The combined inspections of the two operators are compared by one-to-one sticker matching to determine the reproducibility of the operators. This one-to-one sticker matching process is also performed with the user defined search zone coefficient. During the one-to-one sticker matching process, if the center position of another sticker from the other operator's inspection trial of the same casting falls within this zone, then the two stickers are considered to be matched. It is also possible for a sticker to have multiple matching stickers if the center positions of more than one sticker fall within the search zone.

Similar to repeatability, the reproducibility of two operators is also reported as two different percentage values. The first percentage value indicates how well the operators performed in identifying the same supposed defect regions between the two inspection trials of the same casting, reported as percent master cluster match. The second percentage value indicates how well the operators performed in using the same size markings to define the same supposed defect regions between the two inspection trials of the same casting, reported as percent sticker match. The higher the master cluster or sticker percentage values, the better their corresponding reproducibility measurements are. For example, a 0% match would be the lowest, indicating no reproducibility, where a 100% match would be the highest, indicating perfect reproducibility for two operators’ performance on a particular casting.

A visual representation for the reproducibility of two operators can be seen in Figure 5. In this case, operator 1’s trial 1 contains 7 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. Operator 2’s trial 1 contains 5 stickers for 3 supposed defect regions, and trial 2 contains 4 stickers for 3 supposed defect regions for the same casting. Then, the two inspection trials of the operators
are combined to represent the best evaluation of each operator for that casting. Operator 1’s combined trial contains 11 stickers for 4 supposed defect regions, and operator 2’s combined trial contains 9 stickers for 4 supposed defect regions. The two combined inspection trials of the operators for the same casting are then compared. The operators’ two inspection trials for the same casting match 12 stickers out of 20, and 3 master clusters out of 4 in total. For this hypothetical example, the two operators’ reproducibility is 60% for sticker and 75% for master cluster matching, respectively.

Figure 5. Defining reproducibility of two operators with the same casting.
CHAPTER 4. METHODOLOGY

Methodology for this study took place in three stages. The first stage was the collection of the visual quality inspection image data of the castings from three different steel foundries. The next stage dealt with the extraction of the coordinate data from the images collected. The final stage was the analysis of the coordinate data with the use of the analysis methods outlined in Chapter 3.

4.1. Image Data Collection

Image data was collected at three steel foundries, which collectively represent the North American steel casting industry quite well. The companies ranged from 150 to 300 employees, and produced castings for a variety of construction equipment, pump and valve, and industrial equipment applications. One of these companies almost exclusively pours high alloy castings, another almost exclusively carbon and low alloy steel castings, and the third pours about 75% carbon and low alloy with the remainder being high alloy and wear resistant grades.

A similar procedure was used at each foundry for the image data collection. The setup included 2 visual quality inspection operators and 6 castings in Foundry 1, and 10 castings in Foundries 2 and 3. A visual representation of the setup used for the image data collection is displayed in Figure 6.
The castings, chosen randomly by the trial moderator, were marked for defects by the operators employing the same method that they typically use to inspect castings. The only difference was that only one side of the castings was inspected for this study. Each of the two operators inspected each casting twice, on different days, to reduce bias. Between trials, the castings were shot blasted. Figure 7 (a) shows an example casting marked by an operator. After the castings were marked, round stickers of $\frac{3}{8}$ in. in diameter were used to cover the markings as shown in Figure 7 (b).
Figure 7. Casting marked by an operator (a) and stickers applied to the markings (b).
Specific instructions were provided to the trial moderator of each experiment as to how the stickers needed to be placed. The instructions for inspection data collection are included in Appendix A. Careful attention was paid to carry out this process as consistently as possible. Then, digital pictures of the sticker-covered castings were taken.

4.2. Sticker Coordinate Data Extraction

Before the coordinate data extraction could be done, all images went through a series of processes to satisfy specific requirements of the image analysis software, eliminate noise, and assign them operator, trial, and casting numbers in the corners of the images. The image cleaning process was done using Adobe Photoshop. Key steps in the image cleaning process included manually selecting the outlines and the distinguishable characteristics of the castings, and the outlines of the stickers. Repetitive steps such as setting up the images with layers, stroking (coloring) the selected outlines, and formatting the images to the desired size and type were done using automated actions. The final step of the process before saving was to insert identification numbers onto the cleaned images to distinguish the operator, trial, and casting numbers. The main purpose of the image cleaning was to eliminate extraneous image noise. This process did not change the location of the stickers. The cleaned image of the inspected casting shown in Figure 7 is displayed in Figure 8. More samples can also be found in Appendix B. The number on the top left corner corresponds to the operator number. The number on the top right corner corresponds to the trial number. The number on the bottom left corner corresponds to the casting number.
Figure 8. Cleaned image of an inspected casting.

After all the images are cleaned, they are grouped according to the foundry and loaded into DVT FrameWork image analysis software. DVT FrameWork is most commonly used in analyzing images for quality control purposes and is ideal for tasks such as identifying the locations of part features, or in this case, stickers. The inspection parameters and configurations are defined with the use of inspection tools that are drawn on the image very much like a standard drawing program. These inspection tools are called softsensors within the software and are also used to perform specific vision functions. These functions include converting the image into pixel data to further analyze it quantitatively. A sample screenshot of an image data loaded into DVT FrameWork with the softsensors can be seen in Figure 9.
Six different softsensors along with a script were utilized in DVT FrameWork to detect the positions and the radii of the stickers. Then, an ActiveX data link controller was used to output them to a spreadsheet.

Softsensors were utilized to carry out a variety of functions within the DVT FrameWork software. The first step was to recognize the casting and determine its center position and the angle rotation in the image. Next, the operator, trial, and casting numbers from the three corners of the image were detected. Another softsensor was used to detect all the white areas on black backgrounds. This allowed for the detection of casting features along with the stickers. A different softsensor was utilized to separate the stickers from the casting features and determine their center positions along with their areas in pixels. The
stickers' center positions were determined with respect to the casting and adjusted for the orientation and translation of the casting within the image.

A script tool written within DVT FrameWork was used to collect the necessary information from all the sensors, extract the x and y coordinates and the radii of the stickers in pixels, and output that information within a common reference frame to a data link. The full code of the script is included in Appendix C and the logic flowchart is displayed in Figure 10.

Figure 10. DVT FrameWork script logic flowchart.
A DVT FrameWork ActiveX data link controller was then used to extract the sticker data outputted by the script as comma separated values and save them as a MS Excel spreadsheet as the DVT FrameWork inspection process went through all the images collected from a particular foundry. A sample screenshot of the data link controller extracting the sticker coordinate data can be seen in Figure 11.

![Figure 11. DVT FrameWork screenshot with the data link controller extracting sticker coordinate data.](image)

Since each data extraction process used a different casting from a particular foundry, this process needed to be carried out as separate runs to make the necessary adjustments in
the soft sensors used. The saved spreadsheet data consisted of operator number, trial number, casting number, sticker number, x and y coordinates (in pixels), and the radius of the sticker (in pixels) listed for each sticker. A sample screenshot of this data is displayed in Figure 12.

![Figure 12. MS Excel spreadsheet screenshot showing the extracted sticker coordinate data.](image)

4.3. Statistical Analysis via Automated Algorithms

After the sticker coordinate data was saved in a spreadsheet, another MS Excel spreadsheet file with a user interface was created to conduct an analysis. This analysis was conducted via automated algorithms created by macros and was carried out separately for each of the data extraction processes. A sample screenshot of the user interface from the MS Excel spreadsheet used for the analysis can be seen in Figure 13.
The user interface consists of buttons with various functions all connected to macros and a results display area. The buttons are used to organize the functions of the macros, from sorting the data to carrying out the algorithms for analysis.

There are four main macros used to automate the functions of sorting, master cluster, repeatability, and reproducibility. These macros can be run automatically in sequence or individually by themselves. Each of these four macros carries out its operations in a separate sheet so that its results can later be examined separately. There is also a button that can be used to clear all the data and return the interface to its original starting position.

The sorting operation was used to transfer the previously saved sticker coordinate data into the spreadsheet, sort, and label the data for further analysis. A sample screenshot of the sorted data after the sorting operation was done is displayed in Figure 14.
The master cluster operation was used to take the sorted sticker coordinate data and group the stickers into master clusters as discussed in Chapter 3. The default search zone coefficient for the master cluster search was set at 2.5. Although Foundry 1 used the default 2.5 for the search zone coefficient, Foundries 2 and 3 used 3.0 (a 20% increase from the default value) because of their lower image qualities and higher number of stickers that were at an angle to the image view. Further discussion on the determination of search zone coefficients and a sensitivity analysis can be found in Chapter 6. The logic flowchart of the master cluster operation can be seen in Figure 15.
The repeatability operation was used to take the sorted sticker data grouped into master clusters and conduct a one-to-one sticker matching between the different trials of the same castings inspected by the same operator. This process was done for each operator separately and repeated for each casting to determine the repeatability of each operator. The default search zone coefficient for the repeatability search was set at 1.5. Although Foundry 1 used the default 1.5 for the search zone coefficient, Foundries 2 and 3 used 1.8 (a 20% increase from the default value) because of their lower image qualities and higher number of stickers that were at an angle to the image view. The results of the repeatability operation
were then posted on the results display area located at the user interface section. The logic flowchart of repeatability operation is displayed in Figure 16.

![Logic Flowchart](image)

Figure 16. MS Excel repeatability operation logic flowchart.

The reproducibility operation was used to take the sorted sticker data grouped into master clusters and conduct a one-to-one sticker matching between the same trials of the same castings inspected by different operators. This process was done with the combined images of the two trials of the same operator for the same casting and repeated for each casting to determine the reproducibility of the two operators. The default search zone
coefficient for the reproducibility search was set at 1.5. Once again, Foundry 1 used the
default 1.5 for the search zone coefficient, and Foundries 2 and 3 used 1.8 (a 20% increase
from the default value). The results of the reproducibility operation were also posted on the
results display area located at the user interface section. The logic flowchart of
reproducibility operation can be seen in Figure 17.

![Flowchart Image]

Figure 17. MS Excel reproducibility operation logic flowchart.
CHAPTER 5. RESULTS

The analysis of the repeatability and reproducibility results considered two different aspects for both. The first aspect investigated was the variability in detecting the same supposed defect regions, reported as percent master cluster match. The second aspect investigated was the variability in defining the size of the same supposed defect regions, reported as percent sticker match. The higher the percent master cluster or sticker match, the better their corresponding repeatability or reproducibility measurements would be. A 0% match would be the lowest, indicating no repeatability or reproducibility. A 100% match would be the highest, indicating perfect repeatability or reproducibility. Percent master cluster and sticker match were applied for the repeatability and reproducibility measurements for each casting and reported as weighed averages for each foundry.

The results of the analysis showed that there is a significant amount of repeatability and reproducibility error in the visual casting surface inspections. The repeatability of the two operators from the same foundry were found to be very similar. On average, the repeatability measurements were higher than the reproducibility measurements, and the master cluster match percentages were lower than their corresponding sticker match percentages.

The results of the analysis for each foundry are reported by casting in the following tables. The graphical representation of the summary of results for repeatability is displayed separately for each operator in each foundry and is displayed in Figure 18. The summary of results for reproducibility is also displayed separately for each foundry and can be seen in Figure 19.
Table 1. Summary of repeatability results for Foundry 1.

<table>
<thead>
<tr>
<th>Cast #</th>
<th>Operator 1</th>
<th></th>
<th>Operator 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>71%</td>
<td>72%</td>
<td>55%</td>
<td>73%</td>
</tr>
<tr>
<td>4</td>
<td>80%</td>
<td>86%</td>
<td>71%</td>
<td>90%</td>
</tr>
<tr>
<td>5</td>
<td>50%</td>
<td>67%</td>
<td>50%</td>
<td>67%</td>
</tr>
<tr>
<td>6</td>
<td>25%</td>
<td>33%</td>
<td>33%</td>
<td>40%</td>
</tr>
<tr>
<td>7</td>
<td>100%</td>
<td>100%</td>
<td>50%</td>
<td>57%</td>
</tr>
<tr>
<td>Avg</td>
<td>67%</td>
<td>73%</td>
<td>56%</td>
<td>73%</td>
</tr>
</tbody>
</table>

Table 2. Summary of reproducibility results for Foundry 1.

<table>
<thead>
<tr>
<th>Cast #</th>
<th>Operators 1&amp;2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>27%</td>
<td>44%</td>
</tr>
<tr>
<td>3</td>
<td>56%</td>
<td>78%</td>
</tr>
<tr>
<td>4</td>
<td>33%</td>
<td>47%</td>
</tr>
<tr>
<td>5</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>6</td>
<td>40%</td>
<td>55%</td>
</tr>
<tr>
<td>7</td>
<td>50%</td>
<td>77%</td>
</tr>
<tr>
<td>Avg</td>
<td>43%</td>
<td>67%</td>
</tr>
<tr>
<td>Cast #</td>
<td>% MCluster Match</td>
<td>% Sticker Match</td>
</tr>
<tr>
<td>-------</td>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>1</td>
<td>38%</td>
<td>21%</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>53%</td>
<td>53%</td>
</tr>
<tr>
<td>4</td>
<td>57%</td>
<td>61%</td>
</tr>
<tr>
<td>5</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>55%</td>
<td>65%</td>
</tr>
<tr>
<td>7</td>
<td>67%</td>
<td>54%</td>
</tr>
<tr>
<td>8</td>
<td>61%</td>
<td>71%</td>
</tr>
<tr>
<td>9</td>
<td>58%</td>
<td>69%</td>
</tr>
<tr>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Avg</td>
<td>55%</td>
<td>59%</td>
</tr>
</tbody>
</table>
Table 4. Summary of reproducibility results for Foundry 2.

<table>
<thead>
<tr>
<th>Cast #</th>
<th>% MCluster Match</th>
<th>% Sticker Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31%</td>
<td>40%</td>
</tr>
<tr>
<td>2</td>
<td>38%</td>
<td>59%</td>
</tr>
<tr>
<td>3</td>
<td>29%</td>
<td>51%</td>
</tr>
<tr>
<td>4</td>
<td>27%</td>
<td>41%</td>
</tr>
<tr>
<td>5</td>
<td>25%</td>
<td>29%</td>
</tr>
<tr>
<td>6</td>
<td>14%</td>
<td>25%</td>
</tr>
<tr>
<td>7</td>
<td>42%</td>
<td>57%</td>
</tr>
<tr>
<td>8</td>
<td>33%</td>
<td>33%</td>
</tr>
<tr>
<td>9</td>
<td>25%</td>
<td>31%</td>
</tr>
<tr>
<td>10</td>
<td>17%</td>
<td>53%</td>
</tr>
<tr>
<td>Avg</td>
<td>27%</td>
<td>41%</td>
</tr>
</tbody>
</table>
Table 5. Summary of repeatability results for Foundry 3.

<table>
<thead>
<tr>
<th>Cast #</th>
<th>% MCluster Match</th>
<th>% Sticker Match</th>
<th>% MCluster Match</th>
<th>% Sticker Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83%</td>
<td>40%</td>
<td>78%</td>
<td>59%</td>
</tr>
<tr>
<td>2</td>
<td>100%</td>
<td>55%</td>
<td>57%</td>
<td>70%</td>
</tr>
<tr>
<td>3</td>
<td>50%</td>
<td>58%</td>
<td>33%</td>
<td>54%</td>
</tr>
<tr>
<td>4</td>
<td>33%</td>
<td>23%</td>
<td>100%</td>
<td>81%</td>
</tr>
<tr>
<td>5</td>
<td>63%</td>
<td>85%</td>
<td>71%</td>
<td>73%</td>
</tr>
<tr>
<td>6</td>
<td>50%</td>
<td>67%</td>
<td>40%</td>
<td>81%</td>
</tr>
<tr>
<td>7</td>
<td>67%</td>
<td>48%</td>
<td>70%</td>
<td>78%</td>
</tr>
<tr>
<td>8</td>
<td>75%</td>
<td>81%</td>
<td>100%</td>
<td>83%</td>
</tr>
<tr>
<td>9</td>
<td>75%</td>
<td>59%</td>
<td>67%</td>
<td>63%</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
<td>53%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Avg</td>
<td>67%</td>
<td>70%</td>
<td>64%</td>
<td>70%</td>
</tr>
</tbody>
</table>
Table 6. Summary of reproducibility results for Foundry 3.

<table>
<thead>
<tr>
<th>Cast #</th>
<th>% MCluster Match</th>
<th>% Sticker Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>67%</td>
<td>58%</td>
</tr>
<tr>
<td>2</td>
<td>43%</td>
<td>74%</td>
</tr>
<tr>
<td>3</td>
<td>44%</td>
<td>76%</td>
</tr>
<tr>
<td>4</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>5</td>
<td>67%</td>
<td>91%</td>
</tr>
<tr>
<td>6</td>
<td>80%</td>
<td>87%</td>
</tr>
<tr>
<td>7</td>
<td>60%</td>
<td>72%</td>
</tr>
<tr>
<td>8</td>
<td>100%</td>
<td>82%</td>
</tr>
<tr>
<td>9</td>
<td>67%</td>
<td>76%</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
<td>63%</td>
</tr>
<tr>
<td>Avg</td>
<td>65%</td>
<td>82%</td>
</tr>
</tbody>
</table>
Figure 18. Repeatability results for percent master cluster match (a) and percent sticker match (b).
Figure 19. Reproducibility results for percent master cluster match (a) and percent sticker match (b).
CHAPTER 6. DISCUSSION

The discussion section will take a closer look into some of the issues concerning this study. First, the effects of casting cleaning procedures will be looked at. Then, the determination of the search zone coefficients will be explained, followed by a sensitivity analysis. Finally, the implications of the results will be discussed.

6.1. Effects of Casting Cleaning Procedures

There were initial concerns about the effects of the shot blast procedures the castings went through to remove all the stickers and the markings placed by the previous inspection trials. The cause for concern was the possibility of the shot blast operation to change the appearance of the casting surfaces by enlarging some of the surface abnormalities or removing some of the smaller ones. This could have made the defects more or less visible, and therefore, made them easier or harder to detect.

In order to make sure the cleaning procedures did not introduce this unwanted factor into the study, pictures of the castings were taken before and after three shot blast operations for visual comparison. Before and after pictures of a sample casting are displayed in Figure 20. Although, the shot blast operations made the castings look shinier overall, the visual comparison of the pictures showed no obvious change in the surface quality.

The effects of the cleaning procedures were also analyzed quantitatively by looking at the defect detection rate of the operators between their inspection trials. The results of the analysis also showed no significant change in the number of stickers used by the operators between their first and second inspection trials.
Figure 20. Sample casting picture before (a) and after (b) three shot blast operations.
6.2. Determination of Search Zone Coefficients

As mentioned in Chapter 3, the determination of the master clusters, and the one-to-one sticker matching in repeatability and reproducibility operations were all done with the use of user defined search zone coefficients. The search zone coefficient is a number, which, when multiplied by the radius of a sticker, creates a circular search zone around the sticker's center position. If the center position of another sticker from any of the four inspection trials of the same casting fell within this zone in the master cluster operation, then the two stickers were considered to be touching and assigned to the same master cluster. If the center position of another sticker fell within this zone in the repeatability or reproducibility operations, then the two stickers were considered to be matched. It was possible for a sticker to have multiple touching or matching stickers if more than one sticker fell in the search zone, especially from different inspection trials.

The default search zone coefficient for the master cluster operation was set at 2.5. If two stickers were perfectly touching, then their center positions would be exactly 2 radii apart. The extra 0.5 was added to compensate for the distance lost in the borders of the stickers and the imperfections in the image qualities. Similarly, the default search zone coefficient for the repeatability and reproducibility operations were set at 1.5. If a sticker's center position were within the borders of another sticker, then their center positions would be at most 1 radius apart. The extra 0.5 was also added to these operations for the same reasons. A visual representation of these default values can be seen in Figure 21.
Figure 21. Search zone coefficient examples with default values of 2.5 for master cluster (a) and 1.5 for repeatability and reproducibility operations (b).
6.3. Sensitivity Analysis for Search Zone Coefficients

A sensitivity analysis was conducted to examine the effects of varying the size of the search zone coefficients on the total number of master clusters as well as repeatability and reproducibility percentages using the data from Foundry 1. It was observed that a decrease in the size of the search zone coefficient had a greater effect on the results than an increase. It was also noted that as the size of the search zone coefficient increases, the results approach to their maximum or minimum limits. The effects of changing the size of the search zone coefficients on the total number of master clusters of all the castings from Foundry 1 are displayed in Figure 22. Figures 23 and 24 display the effects of the same change on the repeatability and reproducibility percentages of the operators from Foundry 1 respectively.

Figure 22. Effects of search zone coefficients on the total number of master clusters.
Figure 23. Effects of search zone coefficients on the repeatability of operator 1 (a) and operator 2 (b).
6.4. Implications of Results

The results of the analysis showed that there was significant variation in both repeatability and reproducibility measurements from all three foundries. Although the repeatability measurements were somewhat consistent within the foundries, the reproducibility measurements displayed considerably more variation. This poses a particularly big problem in the industry, indicating a need for more operator training and the use of common tools such as comparator plates and work instructions that detail customer requirements.

Another cause for concern was the higher variability detected in the MCluster Match compared with the Sticker Match percentages in the repeatability and reproducibility measurements from all three foundries. The variations in the MCluster Match create a bigger problem for the foundries, as incorrectly identifying or missing a whole defect region is
much worse than identifying an already detected defect region as a bigger or smaller area. This high variability in the defect detection process means that some of the defects go undetected from the inspection process and reach the customer, where some minor imperfections are incorrectly identified as defects, causing unnecessary time and effort lost in the cleaning operations. This uncertainty in the surface defect detection and unpredictability of the defect rates also cause uncontrolled processing times leading to more inefficiencies in cleaning room operations, such as excessive grinding, welding, etc., and adversely effect other aspects such as production scheduling and material handling as well.

All of the implications discussed lead to increased cost of operations at steel foundries. A recent unpublished study by Harwood (2004) found that the costs of cleaning room operations are within the range of $0.76 to $5.34 per square inch (depending on the depth of the defect area) for welding operations, and $0.46 to $1.25 per square inch (depending on the surface quality after welding) for grinding operations. This means that there are high costs associated with every unnecessary defect marking placed by the casting surface inspection operators.
CHAPTER 7. CONCLUSIONS

There are two main outcomes of this study. The first one is the definition of a method to assess measurement error for subjective measurements. The other contribution is the application of this methodology to evaluate the reliability of the visual quality inspection standards used in the steel casting industry.

7.1. Methodology

The methodology created by this study introduces a novel approach to assess measurement error in the visual quality inspection process used in the steel casting industry today. This methodology fills a need in the industry to assess the amount of measurement variability in a subjective, yet important process. The applications of this methodology, however, are not only limited to the steel casting industry. If utilized, it can prove to be useful in other fields, where subjective evaluations are used such as medical image analyses, nondestructive evaluation techniques, and meat carcass inspections.

7.2. Steel Casting Industry

This study also demonstrates the amount of reliability in the visual quality inspection methods used in the steel casting industry. The high amount of variation observed from the results indicates that the current methods are not reliable. This is mostly due to lack of training of the operators. Poor communication between the supervisors and the operators, and lack of leadership from the management also seem to play a role in the results displayed by this study. The absence of a reasonable measurement system for quality has several costly
implications such as uncontrolled processing times, excessive rework, and unacceptable quality standards.
CHAPTER 8. FUTURE WORK

This study introduced a novel approach to assess measurement error in the visual casting surface inspections. However, there are several opportunities where the methodology created in this study can be utilized further. An extension of this study with practical application would be to look into how well the stickers truly characterize the actual defect regions. This can be used to measure the accuracy of operators in correctly identifying the defect regions according to customer specifications.

Another extension with practical application would be to investigate how much does the presence of high variability in the visual quality inspection system of castings cost the steel casting industry. With the help of such a study, cost per variability percentages can be determined (in terms of cost per sticker and/or cost per master cluster) and the amount of savings that can be received from the variability improvements can be predicted.

A theoretical extension of this study would be to take the statistical analysis used to a higher level, possibly applying various models in the analysis of the data in order to extract more information from the results.

Due to the various defects types and their complicated shapes, current casting surface inspections are dependent on the human eye. The presence of the human factor, however, introduces high variability into this process and makes it very subjective. As today’s technology advances to levels, where automation is more common than ever and pilotless planes are becoming a reality, there are opportunities to improve on the automated inspection systems. As advancing technology makes many new tools available to engineers, there is reason to believe that one day the automated systems might replace the human eye.
APPENDIX A. INSTRUCTIONS FOR INSPECTION DATA COLLECTION

1. The inspector will mark the defects on the castings as usual.

2. The trial moderator will cover the markings of the inspector with the stickers.

3. The sticker size to be used will depend on the casting size:
   a. If the longest dimension of the casting is \( \leq 2 \) ft. → use \( \frac{3}{4} \) in. diameter stickers.
   b. If the longest dimension of the casting is \( > 2 \) but \( < 5 \) ft. → use 1 in. diameter stickers.
   c. If the longest dimension of the casting is \( \geq 5 \) ft. → use 1½ in. diameter stickers.

4. More than one sticker may be needed to cover a larger area.
   a. Then the stickers should touch, but not overlap with each other.

5. Finally, pictures of the castings (marked with stickers) need to be taken.

6. The resolution of the digital camera used needs to be set on low to accommodate the size needed for the large number of images.

7. This data collection will be repeated for as many castings as possible up to 10, each casting going through at least 2 different operators twice. If only one operator conducts the test, only the repeatability portion of measurement error can be measured.

8. After each time a casting goes through an inspector, the stickers need to be removed and the original defect markings need to be cleaned through a shot blast process.

9. Being consistent in the data collection process is more important than being accurate.
   The trial moderator may not be able to place the stickers on the most accurate spots, but using consistent technique throughout the inspection trials should lead to good results.

10. After one of the trials, the quality manager should evaluate the inspection markings, to see if additional anomalies should be marked for more processing, or if acceptable anomalies are identified for processing. These observations should be recorded.
Figure A1 displays an example of a casting marked with stickers. If possible, the background needs to be uniform, made up of one solid color. Also, a scaling reference is needed in the background so that the size of the casting can be determined during the analysis.

A marked region covered by a group of stickers will define a defect area like the one shown in Figure A2. Therefore, the defects that fall in the space between the stickers will be covered.

Figure A3 shows an example of a single defect covered by a single sticker.
Figure A4 displays another example of a single defect covered by a single sticker.

Figure A4. Single Line Defect (short).

An example of a single defect covered by two stickers is shown in Figure A5.

Figure A5. Single Line Defect (long).

Figure A6 displays an example of a single defect covered by three stickers.

Figure A6. Single Line Defect (angular).
An example of a single long defect covered by multiple stickers is shown in Figure A7.

Figure A7. Single Line Defect (zigzag).

Figure A8 displays an example of multiple defects covered by multiple stickers.

Figure A8. Multiple Defects.
APPENDIX B. CLEANED IMAGE SAMPLES

Figure B1. Sample cleaned images of inspected casting 4 from foundry 1. Images correspond to operator 1 trial 1 (a), operator 1 trial 2 (b), operator 2 trial 1 (c), and operator 2 trial 2 (d).
Figure B2. Sample cleaned images of inspected casting 4 from foundry 2. Images correspond to operator 1 trial 1 (a), operator 1 trial 2 (b), operator 2 trial 1 (c), and operator 2 trial 2 (d).
Figure B3. Sample cleaned images of inspected casting 4 from foundry 3. Images correspond to operator 1 trial 1 (a), operator 1 trial 2 (b), operator 2 trial 1 (c), and operator 2 trial 2 (d).
class myscript
{

public void inspect()
{

float sticker_angle,dist,x,y,radius;
int i=0,numblobs=Sticker_find.NumBlobs;
String cast_number,operator_number,trial_number,positions="";
byte crlf[]=new byte[2];
crlf[0]=10;
crlf[1]=13;
for(i=0;i<numblobs;i++)
{

pow((Sticker_find.BlobPosition.X[i]-Cast_Locate.Point.X),2));
//DebugPrint("cast x" + Cast_Locate.Point.X + ".");
//DebugPrint("cast y" + Cast_Locate.Point.Y + ".");
//DebugPrint("cast angle" + Cast_Locate.ObjectAngle[0] + ".");
if(Sticker_find.BlobPosition.X[i]==Cast_Locate.Point.X)
{
}
if(Sticker_find.BlobPosition.Y[i]-Cast_Locate.Point.Y>0)
{

sticker_angle=270+Cast_Locate.ObjectAngle[0];
}
{
    sticker_angle = 90 + Cast_Locate.ObjectAngle[0];
}

{
    if(Sticker_find.BlobPosition.X[i] - Cast_Locate.Point.X > 0)
    {
        sticker_angle = 0 + Cast_Locate.ObjectAngle[0];
    }
    if(Sticker_find.BlobPosition.X[i] - Cast_Locate.Point.X < 0)
    {
        sticker_angle = 180 + Cast_Locate.ObjectAngle[0];
    }
}

{

//DebugPrint("blob x " + Sticker_find.BlobPosition.X[i] + ",");

//DebugPrint("cast x " + Cast_Locate.Point.X + ",");

//DebugPrint("cast y " + Cast_Locate.Point.Y + ",");

//DebugPrint("sticker angle top side " + sticker_angle + ",");

}

{

sticker_angle=360-(atan2((Sticker_find.BlobPosition.Y[i] -
(180/3.141592) ) + Cast_Locate.ObjectAngle[0];


//DebugPrint("blob x " + Sticker_find.BlobPosition.X[i] + ",");

//DebugPrint("cast x " + Cast_Locate.Point.X + ",");

//DebugPrint("cast y " + Cast_Locate.Point.Y + ",");

//DebugPrint("sticker angle bottom side " + sticker_angle + ",");

}

if(sticker_angle<0)
{

sticker_angle=sticker_angle+360;

//DebugPrint("< 0 " + i + ",");

}

if(sticker_angle>360)
{

}
sticker_angle = sticker_angle - 360;

//DebugPrint(" > 360 " + i + ", ...");

x = dist * cos(sticker_angle * 3.141592 / 180);
y = dist * sin(sticker_angle * 3.141592 / 180);
operator_number = top_left_reader.String;
trial_number = top_right_reader.String;
cast_number = bottom_left_reader.String;

radius = sqrt(Sticker_find.BlobArea[i] / 3.141592);

positions += operator_number + "," + trial_number + "," + cast_number + "," + (i + 1) + "," + x + "," + y + "," + radius + String(crlf);

DebugPrint(positions);
this.String = positions;
REFERENCES CITED


ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation for my major professor, Dr. Frank Peters, whose guidance and encouragement made this thesis possible. I would like to thank my committee members Dr. Matt Frank and Dr. Iver Anderson as well as Dr. Stephen Vardeman and Malcolm Blair for their time and support.

I would like to express my appreciation for Roy Stevenson, Cory Samuelson, and Brian Harwood for their assistance in data collection, Mike Schmitt for image cleaning, and Nick Wiegand for DVI FrameWork technical support.

I would like to extend my gratitude to Sutee Eamkajornsiri for his expertise and assistance in coding. I would also like to thank the Steel Founders’ Society of America and their member companies which participated in this study.

Finally, I would like to thank the Department of Energy (DOE) Office of Industrial Technology, which provided funding via grant No. DE-FC07-02ID14228. However, any opinions, findings, conclusions, or recommendations expressed herein are those of the author and do not necessarily reflect the views of the DOE.
NAME OF AUTHOR: Gokcer Daricilar

DATE AND PLACE OF BIRTH: JUNE 26, 1980, Istanbul, TURKEY

DEGREES AWARDED:

B.S. in Industrial Engineering, Iowa State University, 2003

MEMBERSHIPS:

Institute of Industrial Engineers, 2003

Alpha Pi Mu Industrial Engineering Honor Society, 2003

Tau Beta Pi Engineering Honor Society, 2003

National Society of Collegiate Scholars, 2000

PROFESSIONAL EXPERIENCES:

Research Assistant, Department of Industrial & Manufacturing Systems Engineering, Iowa State University, 2004

Teaching Assistant, Department of Industrial & Manufacturing Systems Engineering, Iowa State University, 2003